Social and Information Networks

University of Toronto CSC303
Winter/Spring 2021

Week 9: March 15-19 (2021)

This week's agenda

- Mitochondrial Eve
  - Problem setup (Ch 21.7)
  - Wright-Fisher single-parent ancestry model (Ch 21.7)
  - Estimation of time to convergence (Ch 21.8B)
- Bargaining in a Network Exchange Model
  - Power in the network exchange social experiment (Ch 12.1-12.3)
  - Stable outcomes (Ch 12.7)
  - The Ultimatum Game (Ch 12.6)
  - Balanced outcomes (Ch 12.5, 12.8)

Genetic inheritance and networks

Chapter 21 turns its attention to the issue of genetic inheritance, viewed as a random process taking place on a (directed acyclic) network of organisms (species, parts of a genome, etc).

- Motivating example: in 1987, Cann, Stoneking and Wilson published a very striking and to many a very controversial paper
  - Asserted that if one traces their maternal lineage back in time, everyone's lineage traces back to a single woman
  - This woman is called **Mitochondrial Eve**
  - She lived sometime between 100,000 and 200,000 years ago
  - Probably living in Africa
- We'll ignores the issue of the location of Mitochondrial Eve and focuses on the basis (i.e. a model based on various assumptions) for this bold assertion of a common ancestry

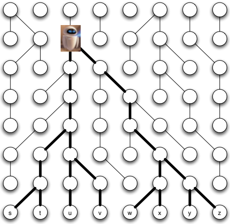
Note: I suggest reading the text as to the caveats about the model (see Ch 21.7)

Modeling the Mitochondrial Eve assertion

- To understand the assertion, we have to make some simplifying biological and mathematical assumptions (see section 21.8 B)
  - The biological assumptions are beyond the scope of the course
    - We will accept them as they are generally accepted to not quantitatively change the conclusions
  - The key biological idea is that "mitochondrial DNA (is to a first approximation) passed on to children entirely from their mothers"
  - The mathematical assumptions do not change any of the conclusions

Mitochondrial Eve continued

- Focussing on mitochondrial DNA, and assuming pure inheritance from the mother, then we can consider a "single parent" ancestry model



[Fig 21.13, E&K]

- The model lets us conclude common mitochondrial DNA ancestry must have originated with a single female Mitochondrial Eve
- The model can also estimate for the time period in which she lived

- This does **not** say that Mitochondrial Eve was the only woman alive at this time, but that our mitochondrial DNA traces back to one woman
- Additionally, our genomic makeup does come from both parents

The Wright-Fisher single parent ancestry model

- The Wright-Fisher model not only applies to mitochondrial lineage, but also to general asexual reproduction
- Additional simplifying assumption for tractability:
  - assume generations are synchronized
  - assume a fixed population of N individuals throughout the entire period of time
- Inconsistent with the fact that world population is growing
- Ultimately does not change the nature of the conclusions or even the nature of the analysis
  - In fact, once we accept that populations are growing, it is clear that certain individuals must be having multiple children which is also part of the model

Single parent ancestry model continued

- Assume that generations are completely synchronized:
  - the generation of N individuals at time t give rise to the next generation of N individuals at time $t + 1$.
- Each individual at time t+1 has its "single parent" chosen uniformly at random from the previous generation
  - A significant assumption given geography, ethnicity, etc...
  - To reconcile this (with respect to the assertion of a single Mitochondrial Eve), we need to understand the extent to which individual communities can be isolated
    - ★ Ultimately, the timing for when common ancestry would have taken place is not impacted by this assumption

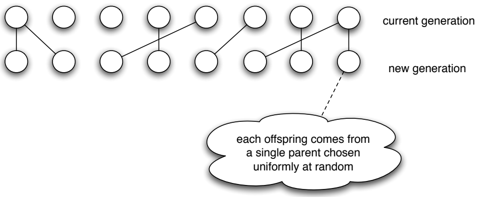
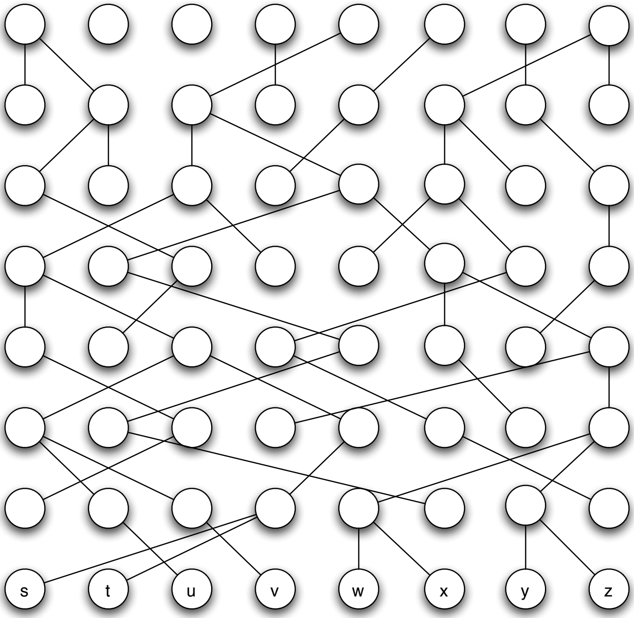
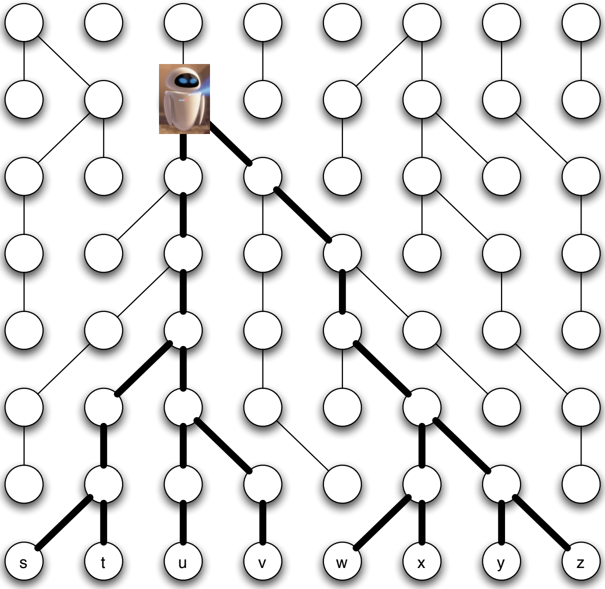


Figure: [Fig 21.11, E&K]

More generations of the model



Ancestry depicted.



The analysis for estimating the time that the model coalesces on Mitochondrial Eve

Section 21.8B provides a mathematical analysis for estimating the time when a common ancestor (in the single parent model) will be reached. Along the way, some simplifying mathematical assumptions are made but these assumptions are easily defended and are not of the same nature as biological assumptions.

Suppose we have a total population of N and at some point of time $t+1$ that we are down to k candidates (lineages) for a common ancestor. We want to consider the probability that two lineages will collide so that there be (at most) $k-1$ candidates.

The analysis

Instead of moving time forwards until the current generation shares an ancestor, we will move time *backwards* until a common ancestor emerges. We will start by considering k nodes, and seeing how probable it is that they do not have the same parent parent.

Case: $k = 2$. Say the active lineage is individuals $\{a, b\}$. Then the probability that b does not share a's parent is $1 - \frac{1}{N}$.

Case: $k > 2$. Lets consider the probability that none of the k nodes share a parent. There will be no collapsing if the second node doesn't collide with the first, the third doesn't collide with the first two, etc, so this means that the probability of no collapsing is :

$$(1 - \frac{1}{N})(1 - \frac{2}{N}) \cdots (1 - \frac{k-1}{N})$$

The analysis continued

The previous product

$$(1 - \frac{1}{N})(1 - \frac{2}{N}) \cdots (1 - \frac{k-1}{N})$$

is at most:

$$1 - \left(\frac{1 + 2 + \cdots + k - 1}{N} \right) + \frac{g(k)}{N^2}$$

where $g(k)$ depends only on k and not on N.

For any fixed k, the latter term is relatively negligible and we can say that the probability that none of the k share a parent is $1 - \frac{k(k-1)}{2N}$.

The analysis continued

Fact: If we have a binary random variable Y_k (i.e., a heads coin flip) that is true with probability p, then the expected number of independent samples until Y_k is true (denoted $E[X_k]$) is exactly $1/p$

- if the probability is at least p, then the expected time can only be shorter.

Look familiar? Remember the geometric distribution, and the decentralized search tutorial! We're going to do (basically) the same proof ;)

Therefore, letting X_k denote the time to collapse from k to less than k lineages, then $E[X_k]$ is approximated by $\frac{2N}{k(k-1)}$

Note: Initially when k is large, the decrease is expected every generation going back. But when k is a small constant, then the expected number of generations to show a decrease is proportional to N.

Depiction of the lineages colliding

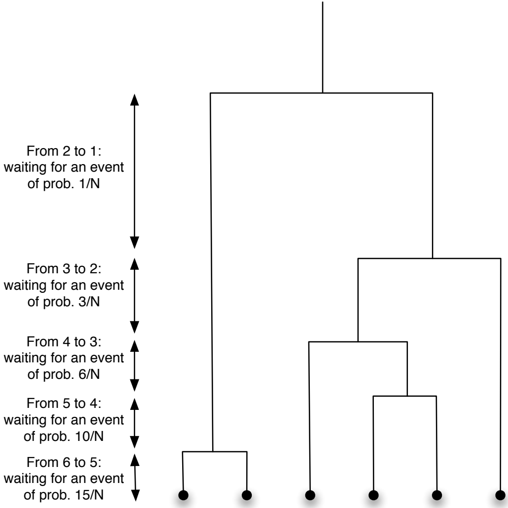


Figure: Assuming no three lineages collide simultaneously. [Fig 21.1(a), E&K]

Finishing the analysis

Let $X^k = X_k + X_{k-1} + \cdots + X_2$ be the number of generation to reach a common ancestor starting from a lineage of k individuals.

Note: To simplify the analysis we are assuming that k lineages will always collapse into $k-1$ lineages. This assumption is wrong, but provides a good estimate.

Since $\mathbb{E}[X_j] = \frac{2N}{j(j-1)}$ and $\frac{1}{j(j-1)} = \frac{1}{j-1} - \frac{1}{j}$, by linearity of expectations we have:

$$\begin{aligned}
\mathbb{E}[X^k] &= \sum_{j=1}^{k} \frac{2N}{j(j-1)} \\
&= 2N\left(\left[\frac{1}{1} - \frac{1}{2}\right] + \left[\frac{1}{2} - \frac{1}{3}\right] + \cdots + \left[\frac{1}{k-1} - \frac{1}{k}\right]\right) \\
&= 2N\left(1 - \frac{1}{k}\right)
\end{aligned}$$

Note: Further more detailed analysis is consistent with the basic analysis that was presented in the text.

Recap

- Mitochondrial Eve
  - Problem setup (Ch 21.7)
  - Wright-Fisher single-parent ancestry model (Ch 21.7)
  - Estimation of time to convergence (Ch 21.8B)

Chapter 12: Bargaining and Power in Networks

We begin a subtle and fascinating topic, namely how individuals in a network come to agreement on an outcome. This chapter is part of a larger subject called cooperative game theory and to some extent touches on behavioural game theory. As previously discussed, we have a course (CSC304) which covers game theory and in our course we will only present what is necessary regarding game theory. What we need is rather minimal (e.g., as when we were discussing network coordination in chapter 19).

But perhaps here is a good place to mention some basic game theory concepts to keep in mind (and again we have at least implicitly seen these concepts in our discussions to date). The following is a very brief set of informal comments.

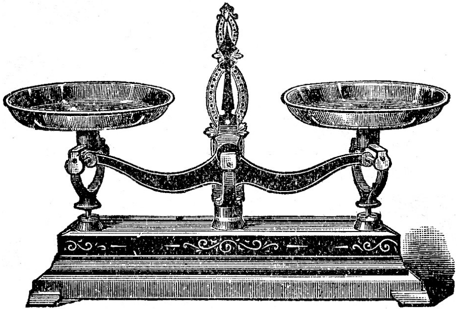
A few more comments on game theory concepts

- Individuals (agents) have strategies or actions and employ a (pure or mixed/randomized) strategy so as to act in self interest, always trying to maximize benefit or minimize cost.
- **Note:** There is a lot of subtlety in benefits and costs
  - often cannot be explained simply in monetary terms (or one must assign monetary values to subjective values)
- Agents are acting in self interest implies that their actions are decentralized
  - *Mechanism design* concerns how a central agent can introduce incentives to influence agents
  - An example of a result in Mechanism Design is Gibbard-Satterthwaite theorem, which states that any voting rule is either
    - ★ Dictatorial
    - ★ Only selects the winner from a set of two candidates
    - ★ Is susceptible to tactical voting

Game theory concepts: Equilibrium

Definition (Equilibrium)

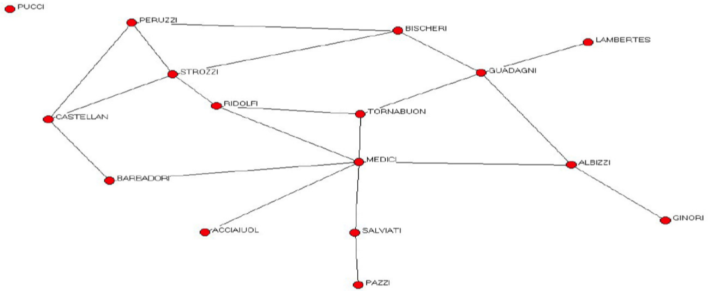
A state in which no agent has an incentive to change their strategy assuming no one else is changing



- Appeared in Schelling segregation model in Chapter 4, structural balance in Chapter 5, and will be important in Chapter 12 and the study of relative power
  - we will see them again in stable matchings and traffic equilibria

Power as a relative relation between people

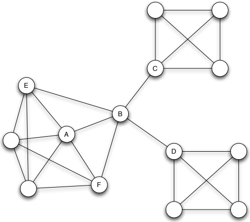
- Power between individuals can come from two distinct sources:
  - The pivotal position of the person in the network.

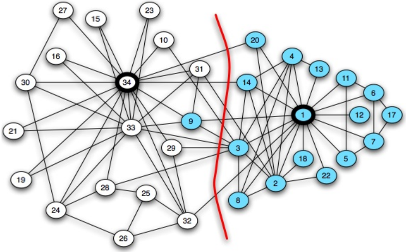


    - ⋆ In the first week we mentioned the network of Florentine marriages and how the *centrality* of the Medici family was said to have conferred power to the Medicis
    - ⋆ In the second week of the course we discussed the *bridging capital* and the *bonding capital* of a node
  - The relative reputation, status, official position, exceptional attributes (intelligence, finances), etc.

Power: Bridging and bonding capital of nodes

The early chapters of the text provided some insights about the importance of centrality and bonding capital and bridging capital with regard to the *flow of information* and *trust*.



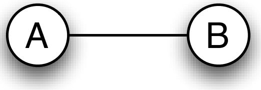


Power arising from asymmetries in pairwise relations

- In contrast, Chapter 12 considers **power** in terms of the relationship between two individuals that results in different division of value *bargaining network*
  - The imbalance in assigned values corresponds to the imbalance in their relative power
- **Note:** In this context, centrality can sometimes be misleading.
- The above is an informal definition of power, but the study of power in the context of imbalance is a well studied concept with precise definitions
- We will isolate power due to position in a network, and ignore the status aspects
- For motivation we begin with some illustrative network examples, we will follow this with a social experiment that will provide insight, and will in turn lead to precise definitions

Some illustrative examples

- Assume \$1 is placed on each edge of the network
  - each node trying to reach an agreement (within a fixed amount of time) on how to split the dollar
  - each node can only deal with at most one other adjacent node
    - In graph theoretic terms, this pairing of nodes is a *matching*: a subset of edges such that no node is adjacent to more than one edge in the matching
- Who will have relative power (i.e., receive more than half a dollar in the following networks)?

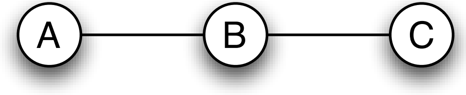


(a) *2-Node Path*

Does either party have an advantage?

No; a $\frac{1}{2} - \frac{1}{2}$ split is a reasonable predicted split that is observed in the experiments.

A three node path



(b) *3-Node Path*

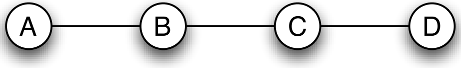
What matching might occur and who each holds power?

Clearly since we need a matching, either A and C will have to be left out. Intuitively then, node B holds much more power than A or C. The basic theory and experiments support this intuition.

What fraction of the $ would you expect B to obtain in negotiating between A and C?

There is a difference between the basic theory and the social experiments. In the experiments , B gets a $(\frac{5}{6})^{th}$ fraction of the $. The basic theory would predict that B gets all almost all of the $. Why the difference?

A four node path

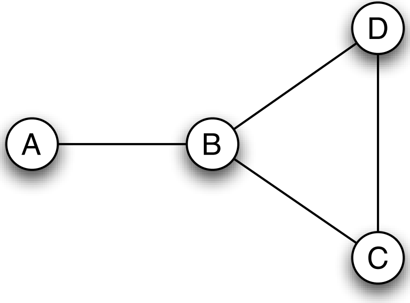


(c) *4-Node Path*

What matching might occur and how might the money be split? Would B get more or less in this four node network than in the previous three node path?

Here the experiments show that B gets a fraction of between $\frac{7}{12}^{th}$ and $\frac{2}{3}^{rd}$ of the \$, less than what we obtained in the three node network. Why?

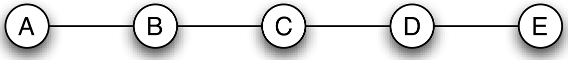
The stem graph in figure 12.3



What matching might occur and how might the money be split? Would B get more or less in this stem network than in the previous three and four node paths?

Experiments show that B in the stem graph makes slightly more money than B in the four node path (but less than in the three node path). Why?

A five node path



(d) *5-Node Path*

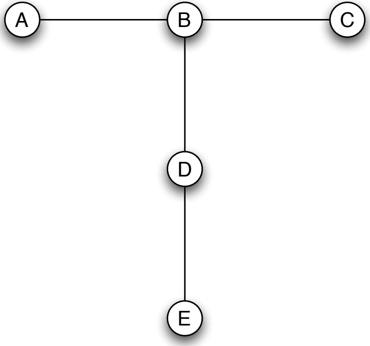
Does C have any power (i.e. fraction of money obtained) compared to other nodes?

Intuitively B and D have most of the power in the five node path network. The text states that in experiments, C has slightly more power than A or E.

Note that C is the most central node in terms of being on all shortest paths. However, this has not translated into substantial batgaining power.

Another graph to consider

The previous examples may help us reason about the following example from the text.



The *network exchange* social experiment

The following network exchange social experiment (and variants) is repeated a number of *rounds* so that some form of learning is taking place. There are many variants and the text presents one particular setting.

- Individuals (not knowing each other since we want to focus on the network aspects and not on the status, etc. of individuals) are placed at computer terminals and can interact with certain other individuals.
- In a complete information setting, one might see the entire network. The text considers the setting where an individual only knows and negotiates with their neighbouring nodes.
- For some known duration on time for a given round, negotiations take place for sharing say one $ on each edge. (We could allow larger and different sums for each edge). Once a pair have decided how to share the $, they leave the game.
- There is one more important condition on the experiment; namely in any given round, the outcome has to be a matching. This is called the *1-exchange rule*.

How much do these experimental findings depend on the exact setting.

We would, of course, like to have results that are robust and do not differ that much in the exact "details".

- Results are reasonably robust with regard to how much network information is available
- Results are consistent across different countries and different cultures

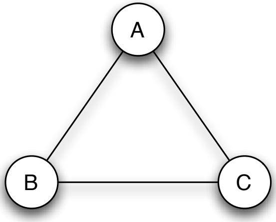
Question: What are we not robust to?

- The 1-exchange rule is a definite factor impacting the results
  - In certain networks, substantially different findings result if individuals can negotiate two or more exchanges in a round
  - In graph theory terms this is a b-matching; nodes can be adjacent to up to b edges in the matching
- Anonymity is important
  - Higher status individuals tend to inflate their "options", and those of lower status tend to underplay their options

Do all experiments converge in a consistent manner?

In simple networks, each round tends to come to consistent outcomes within the specified time limits.

However, there are networks where this is not the case. Consider the following triangle graph:



Question: Notice anything?

Any two of the nodes can wind up excluding the other. Hence we would expect that the final outcome in any round will be determined by the two nodes who get to settle just before the time deadline.

A mathematical perspective: The Nash Bargaining Solution

We would like to model the pure 1-exchange experiments (with anonymous participants)

First, we would like to understand which outcomes will be *stable*. Without having a stable outcome, we cannot hope for participants to converge in any consistent way.

Conversely, we would expect that over enough rounds, participants would learn to converge to a stable outcome. Stable outcomes are equilibria and like most games, there can be many stable outcomes for a network exchange process.

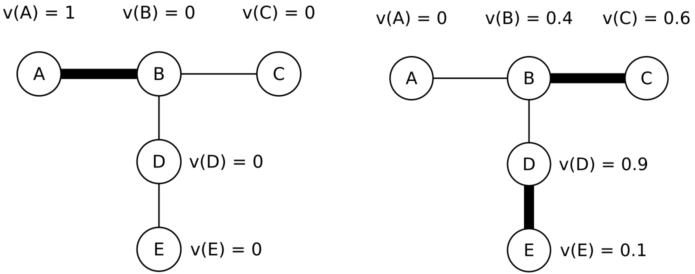
John Nash (the same Nash who showed that all finite games have mixed equilibria) introduced a specific stable outcome, the *Nash Bargaining Solution*.

Outcomes

We first define an outcome for the network exchange social experiment where every edge is worth 1$.

An outcome in a network exchange process on a graph $G = (V, E)$ is a pair (M, v) where $M \subseteq E$ is a matching and the value function $v : V \to [0, 1]$ satisfies:

- For every edge $e = (x, y) \in M$, $v_x + v_y = 1$.
- If a node $x \in V$ is not part of the matching M (i.e. does not appear as a vertex in any edge $(x, y) \in M$), then $v_x = 0$.



Stable outcomes

In a stable outcome, no agent (i.e. node) x can propose to an adjacent agent y, an offer that would improve both of their current outcomes.

Stable Outcomes

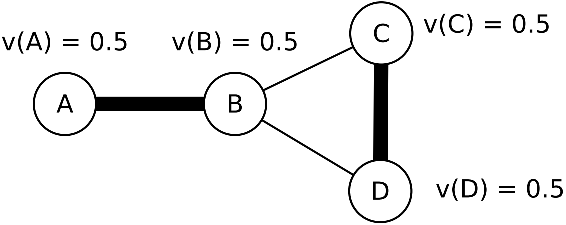
An outcome (M, v) for a network exchange process is *stable* if for every edge $e = (x', y') \in E \setminus M$, $v_{x'} + v_{y'} \geq 1$.

Since we are assuming that each edge has exactly one \$ on each edge, clearly $v_x + v_y = 1$ for each edge $(x, y) \in M$, the matching.

Suppose $v_{x'} + v_{y'} < 1$ for an edge $(x', y') \notin M$. Then the matching is unstable as there is a surplus of $s = 1 - v_{x'} - v_{y'}$ that can be shared between x' and y' and there is no reason for them not to share this surplus and increase both their values.

Which stable outcome?

Stable solutions are necessary but there can be many stable solutions and some are more natural (in the sense of corresponding to real behaviour) than others.



Which stable outcome?

Suppose $(x, y) \in M$. What if x and y have other options other than to be in a given matching? Suppose that x (respectively, y) has an "outside option" of o_x (resp. o_y). Then $o_x + o_y \leq 1$ or else (x, y) could not be in a stable matching as either x or y would be better off taking their outside option.

The Nash bargaining solution would be to keep (x, y) in the matching and equally divide up any surplus from the outside options. That is, if $s = 1 - o_x - o_y$, then set $v_x = o_x + \frac{s}{2} = \frac{o_x + 1 - o_y}{2}$ and $v_y = o_y + \frac{s}{2} = \frac{o_y + 1 - o_x}{2}$. And hence we get:
$v_x + v_y = 1$ with (x, y) in the matching.

Why extreme outcomes are not real outcomes

As stated earlier in this chapter, in the three node path example, the theory thus far would predict that B will obtain the entire \$. But we are told that in experiments, more typically B gets a fraction $\frac{5}{6}$ and one other node gets a fraction $\frac{1}{6}$.

This can be explained once we understand that individuals (i.e., real people) are not driven solely by monetary payments. The "real value" to an individual may include some notion of fairness, pride, etc. When we consider these factors, we can see why in these experiments, extreme solutions (which sometimes are the only theoretically stable solutions) are not the actual outcome.

In the following ultimatum game, we can perhaps better understand why participants tend to think beyond monetary rewards.

Another network exchange game: the so-called "Ultimatum Game"

We again are considering how two individuals divide a \$. But now we have the following experiment:

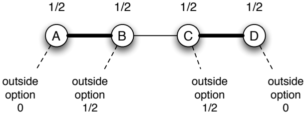
- One person (say A) is given one \$ and is told to propose a division of it to person B.
- Person B is then given the option of accepting the share offered or rejecting the offer.
- If B accepts, the game is over with the division as given by A. If B refuses then each person **gets nothing**.
  Aside: This is a little like the "I cut-you choose 2-person cake cutting algorithm" which insure "fairness".

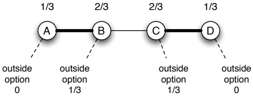
This is a one-shot experiment between people who do not know each other. What do we expect to happen?

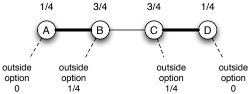
Now in strictly monetary terms, person B should accept any offer (even a \$.01). But this is not what happens in experiments. In experiments, A tends to offer B about one third of the \$. Why?

Not all stable outcomes are "natural"

As we stated, there can be many stable outcomes for a given network. But some do not appear as natural as others and, in particular, stable outcomes can be "extreme solutions" that do not represent what we believe to be more realistic. Which of the following stable outcomes might be more expected "in practice"?







Balanced outcomes

It turns out that the $\frac{1}{3}, \frac{2}{3}$ split between A and B and also between C and D is what happens more in experiments and can be considered "more natural" in the following way.

The equal $\frac{1}{2}$ split amongst all parties does not reflect the relative much better bargaining position of B and C. In contrast, the $\frac{1}{4}, \frac{3}{4}$ split between A and B and also between C and D, seems to be giving B and C too much power given what we have been saying about how humans behave when taking say fairness, pride, etc into account.

Can we give a mathematical explanation for why the $\frac{1}{3}, \frac{2}{3}$ split should be a likely outcome?

It turns out that the $\frac{1}{3}, \frac{2}{3}$ split is the Nash Bargaining solution which we argued seemed like a fair way to divide up surpluses.

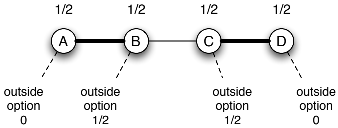
What is a balanced outcome?

Balanced outcomes

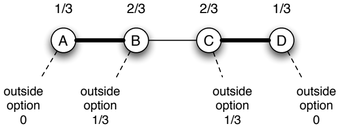
An outcome (M, v) is balanced if for every edge in the matching M, the split of money $\{v_x\}$ is the Nash bargaining solution for each node x, given the (best) outside options for each node.

Fact: For every exchange network with a stable outcome, there exists a balanced outcome.

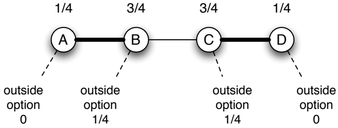
Balanced and unbalanced outcomes for the four node path



(a) *Not a balanced outcome*



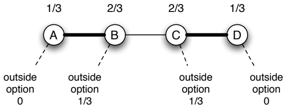
(b) *A balanced outcome*



(c) *Not a balanced outcome*

Checking that the balanced outcome is the Nash Bargaining solution

Let's check that the balanced outcome is indeed the Nash Bargaining solution.



Why is the best outside option for B (and similarly for C) equal to $\frac{1}{3}$?

B has the option of offering $\frac{1}{3}$ (or maybe $\frac{1}{3} + \epsilon$ for some small $\epsilon > 0$) to entice C to leave its current match with D. Of course, A has no outside option so we we can calculate that surplus for the matched edge (A, B) is $s = 1 - o_A - o_B = \frac{2}{3}$ and hence the Nash bargaining solution would be:

- $v_A = o_A + \frac{s}{2} = 0 + \frac{1}{3} = \frac{1}{3}$
- $v_B = o_B + \frac{s}{2} = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$

  which is consistent with the balanced outcome.

Similarly, C and D follow the Nash Bargaining solution.

Recap

- Bargaining in a Network Exchange Model
  - Power in the network exchange social experiment (Ch 12.1-12.3)
  - Stable outcomes (Ch 12.7)
  - The Ultimatum Game (Ch 12.6)
  - Balanced outcomes (Ch 12.5, 12.8)